

OAI-PMH and the Peer-Review Process

Marko Antonio Rodriguez

Digital Library Research & Prototyping Team
Los Alamos National Laboratory

&

Center for Evolution, Complexity, and Cognition
Vrije Universiteit Brussel

&

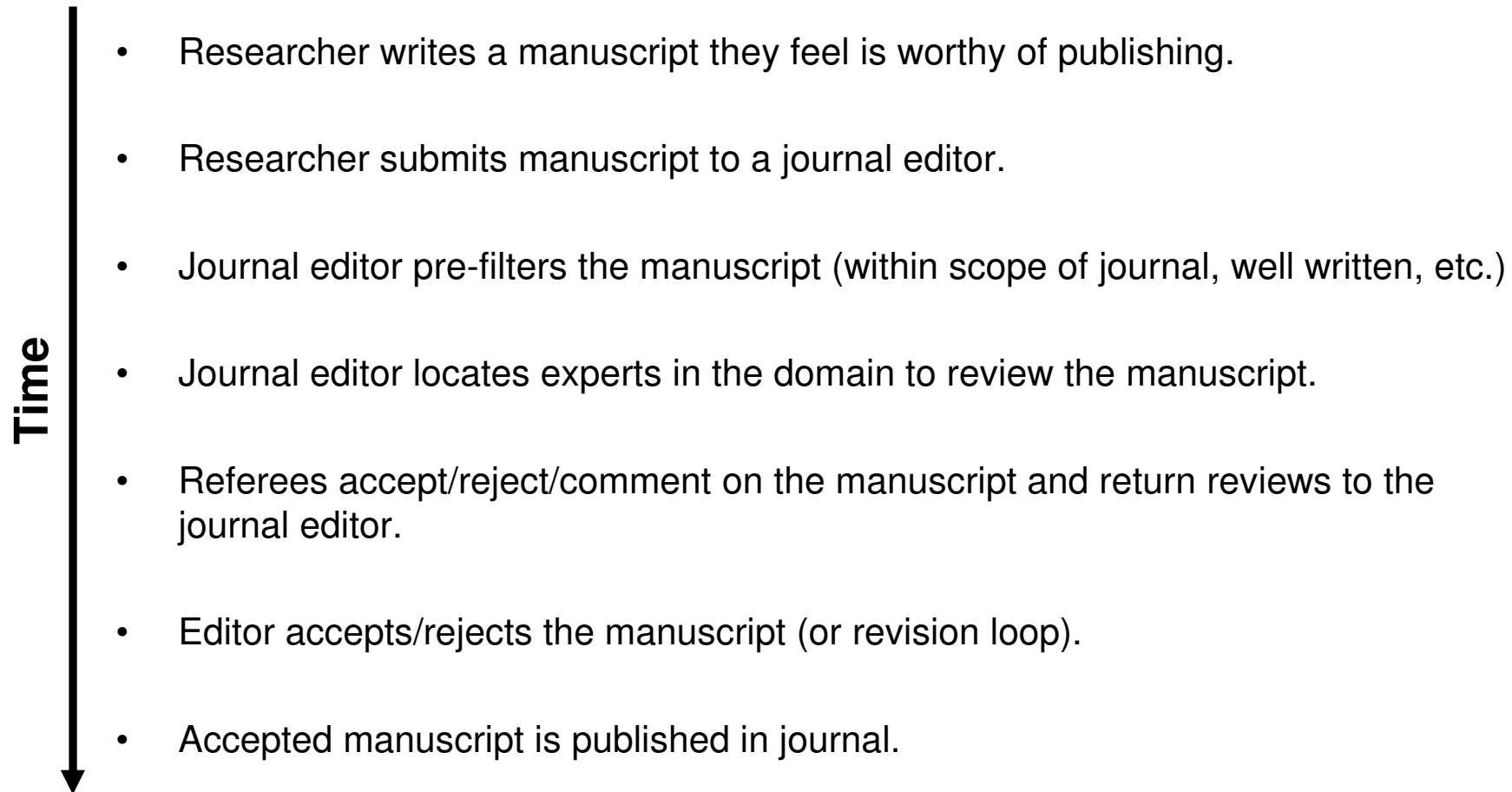
Computer Science Department
University of California at Santa Cruz

email: marko@lanl.gov . URL: <http://www.soe.ucsc.edu/~okram>

Purpose of this Talk

- Describe a peer-review model that revolves around OAI repositories.
- The model removes the need for editors and publishers in scholarly communication.
- The only human components are authors and referees.
- The model can be implemented as a OAI service-provider.
- The peer-review service is able to solicit referees, aggregate referee evaluations, and generate peer-review metadata for the resource's metadata record. (the editor's role)
- The OAI repository provides the information dissemination infrastructure. (the publisher's role)

Overview of the Current Peer-Review Model



Overview of the Proposed OAI Peer-Review Model

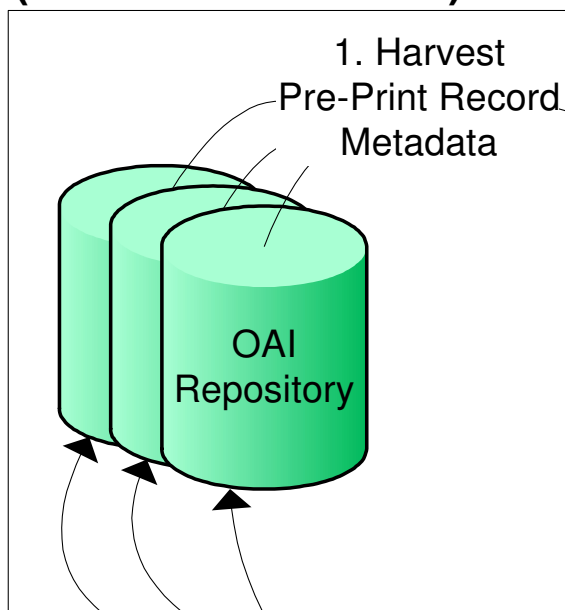
Time



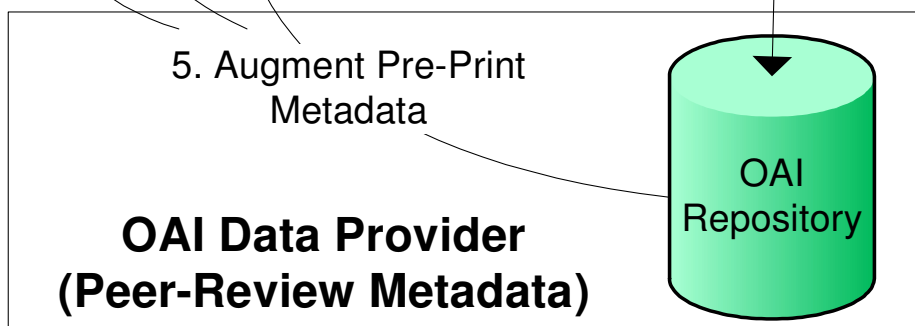
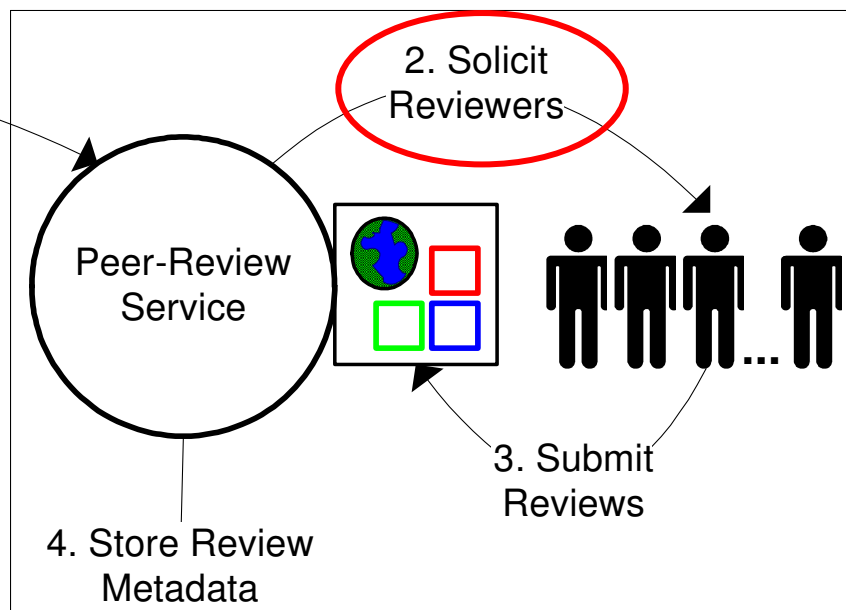
- Individual writes a manuscript they feel is worthy of publishing.
- Individual submits manuscript to OAI repository.
- Peer-review service-provider harvests those e-manuscripts that are worthy of review (i.e. high usage stats, high citation stats, no Journal-Ref, within a certain ACM classification, author requested, community requested, etc.).
- Peer-review service locates experts in the domain to review the manuscript.
- Referees review the manuscript and provide an evaluation by way of an online interface.
- Peer-review service aggregates referee scores and generates the manuscripts peer-review metadata.
- OAI repository provides the manuscript and its associated peer-review metadata to the public.

Overview of the OAI Peer-Review Architecture

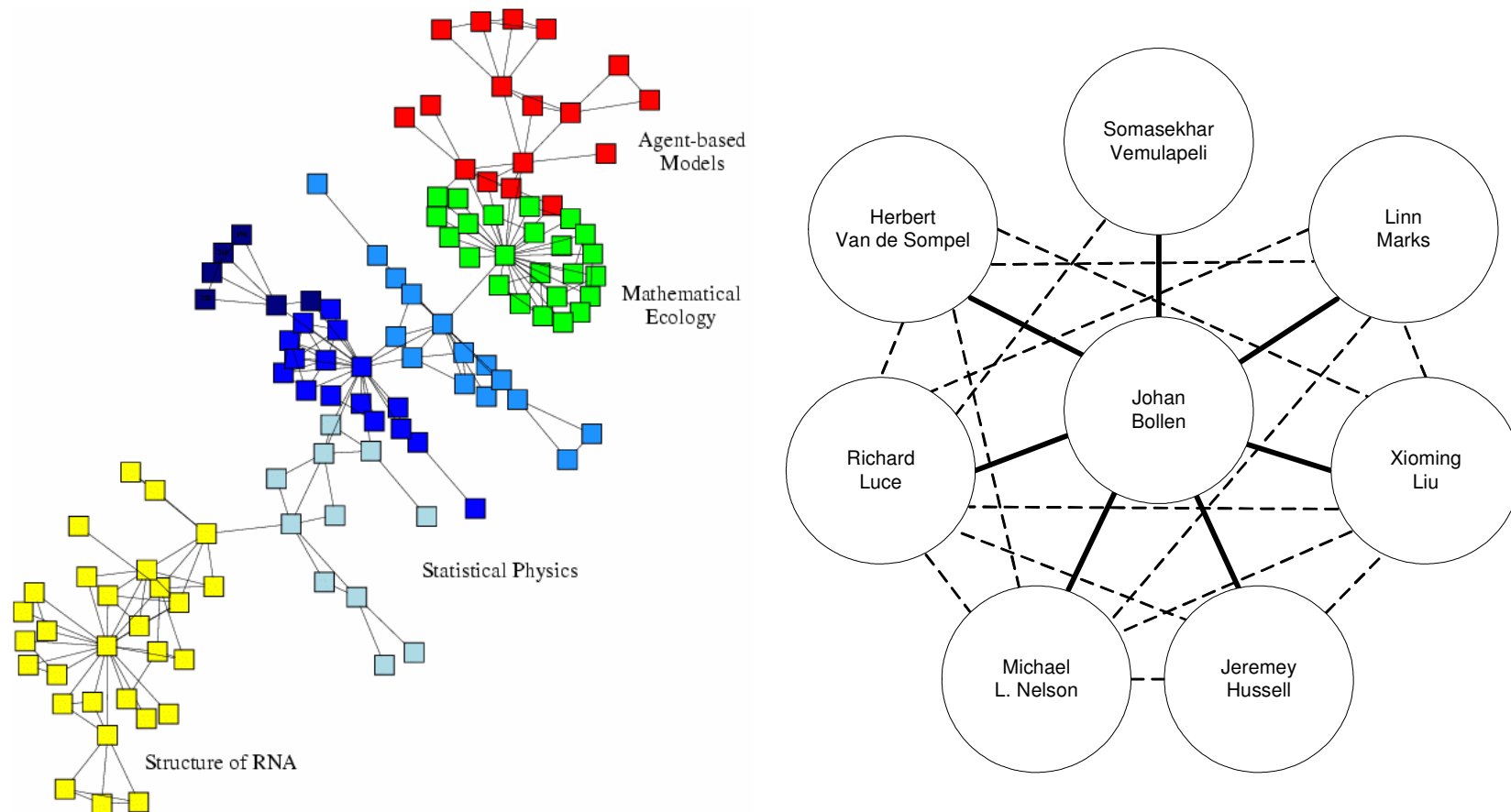
OAI Data Provider (Pre-Print Metadata)



OAI Service Provider

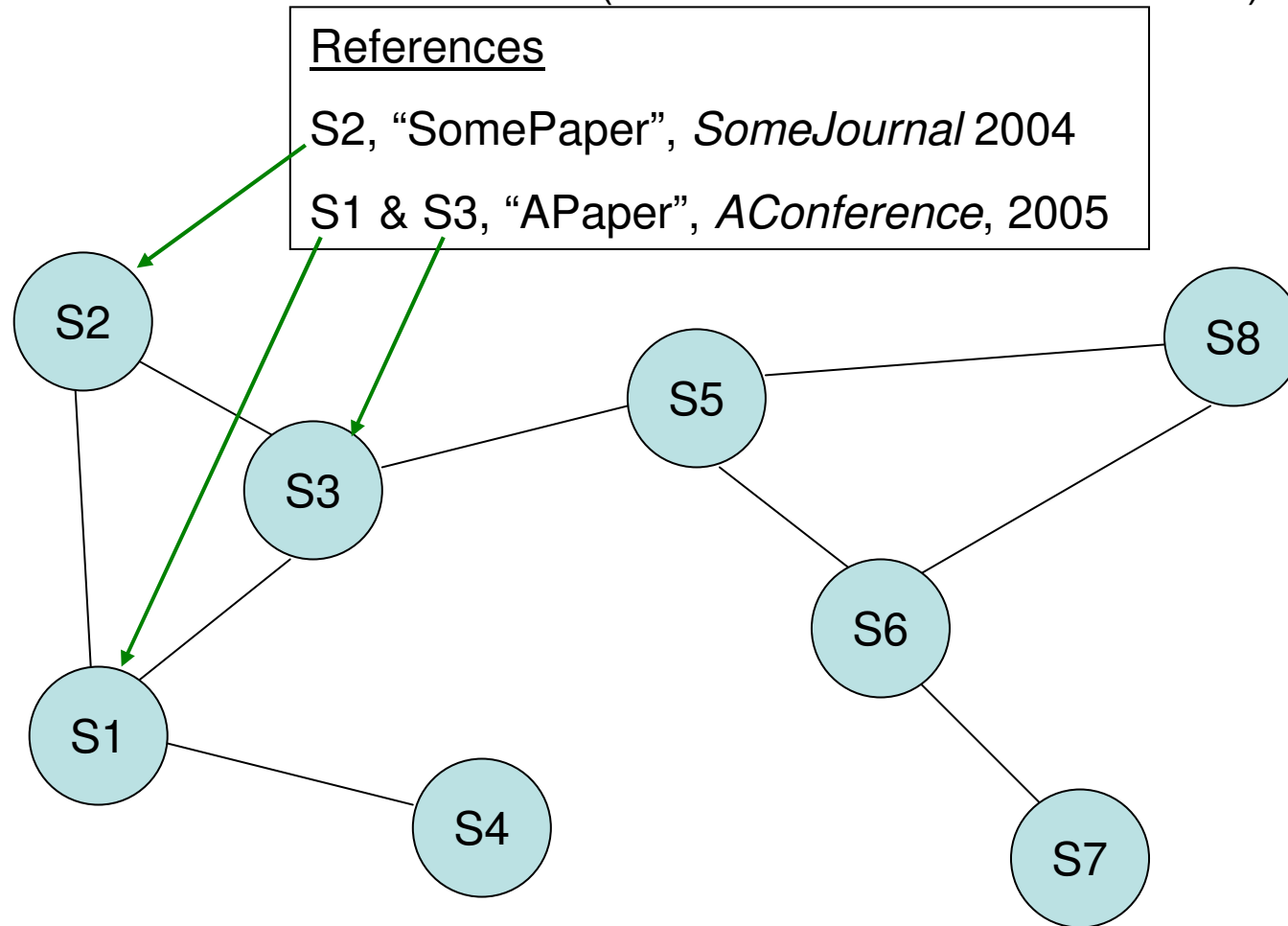


Co-Authorship Networks as a Model of Expertise



Automatic Solicitation of Referees

Harvested Pre-Print (or citation metadata if available)



DEMO

- <http://127.0.0.1:8080/peerper/>

Referee Influence for this talks associated Manuscript

Rodriguez, M.A., Bollen, J., Van de Sompel, H.,

"The Convergence of Digital-Libraries and the Peer-Review Process", *Journal of Information Science [in press]*, September 2005.

Referee Name	Influence	Recent Interests Related to Paper
Sompel, HV	0.09844	OAI-PMH and Co-Authorship Networks
Bollen, J.	0.08594	Digital-Libraries and Network-Based Impact Metrics
Carr, L.	0.08516	Digital-Libraries and Open Archive Services
Hall, W.	0.08066	Knowledge Management and Digital-Libraries
Rocha, L.M.	0.07892	Document Recommendation Systems
Lagoze, C.	0.05328	OAI-PMH and Digital-Library Architectures
Harnad, S.	0.04883	Open Citation Linking and Digital-Library Architectures
Hitchcock, S.	0.04177	Electronic Journals and Citation Linking
Blake, M.	0.04156	OAI Repositories and Citation Linking
Jiao, Z.	0.03386	E-Print Services
Bergmark, D.	0.03262	Digital-Libraries and OAI-PMH
Miles-Board, T.	0.02049	Digital-Libraries
Davis, H.C.	0.01211	Digital-Libraries and Adaptive Linking
Roure, D.D.	0.01125	Dissemination of Scientific Information Services
French, J.C.	0.01081	Digital-Library Distributed Searching and Interfaces
Bailey, C.	0.01043	Digital-Libraries and Distributed Media
Brody, T.	0.00986	CiteBase and Open Citation Linking

Using Real Peer-Review Bid Data to Validate Algorithm

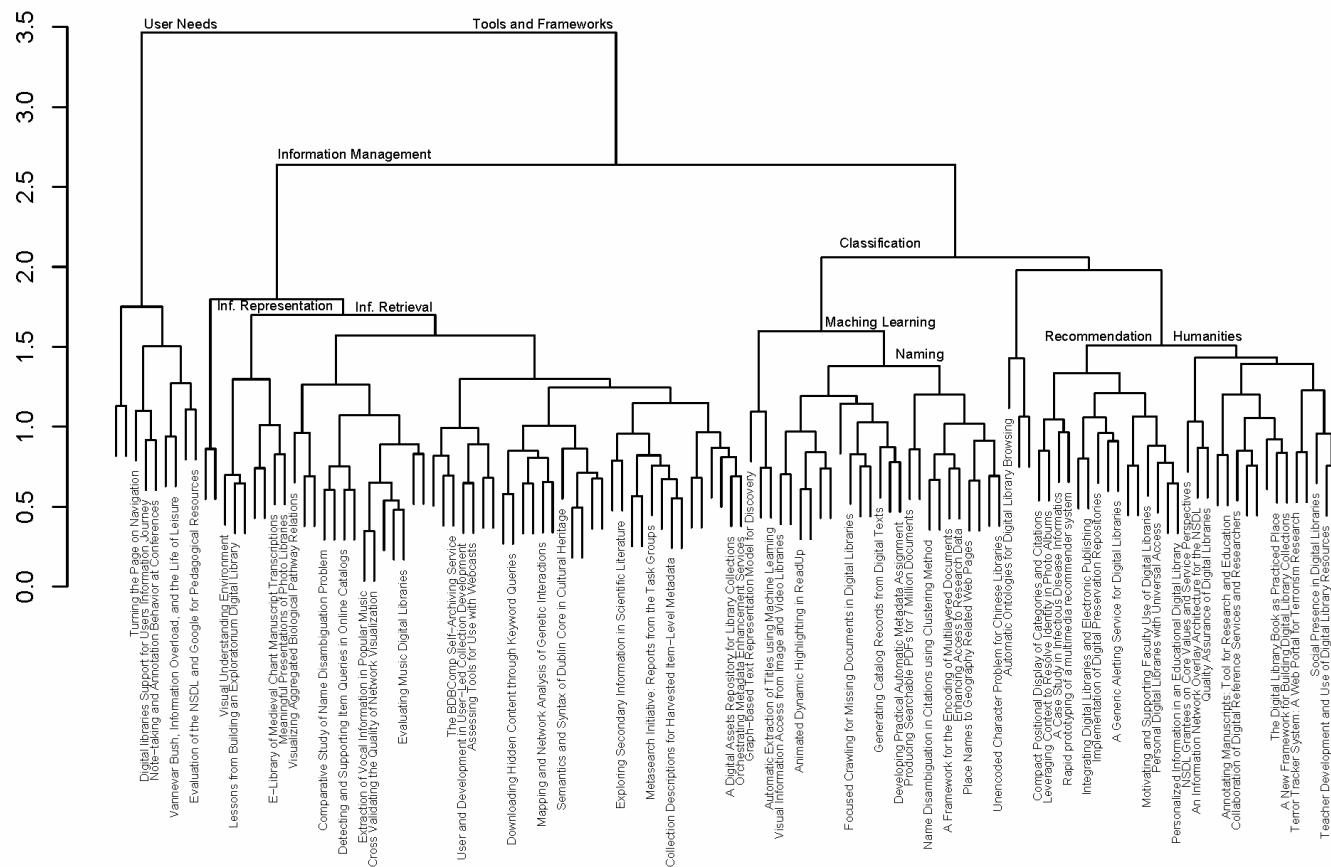
- Received Bid Data and Submission Archive

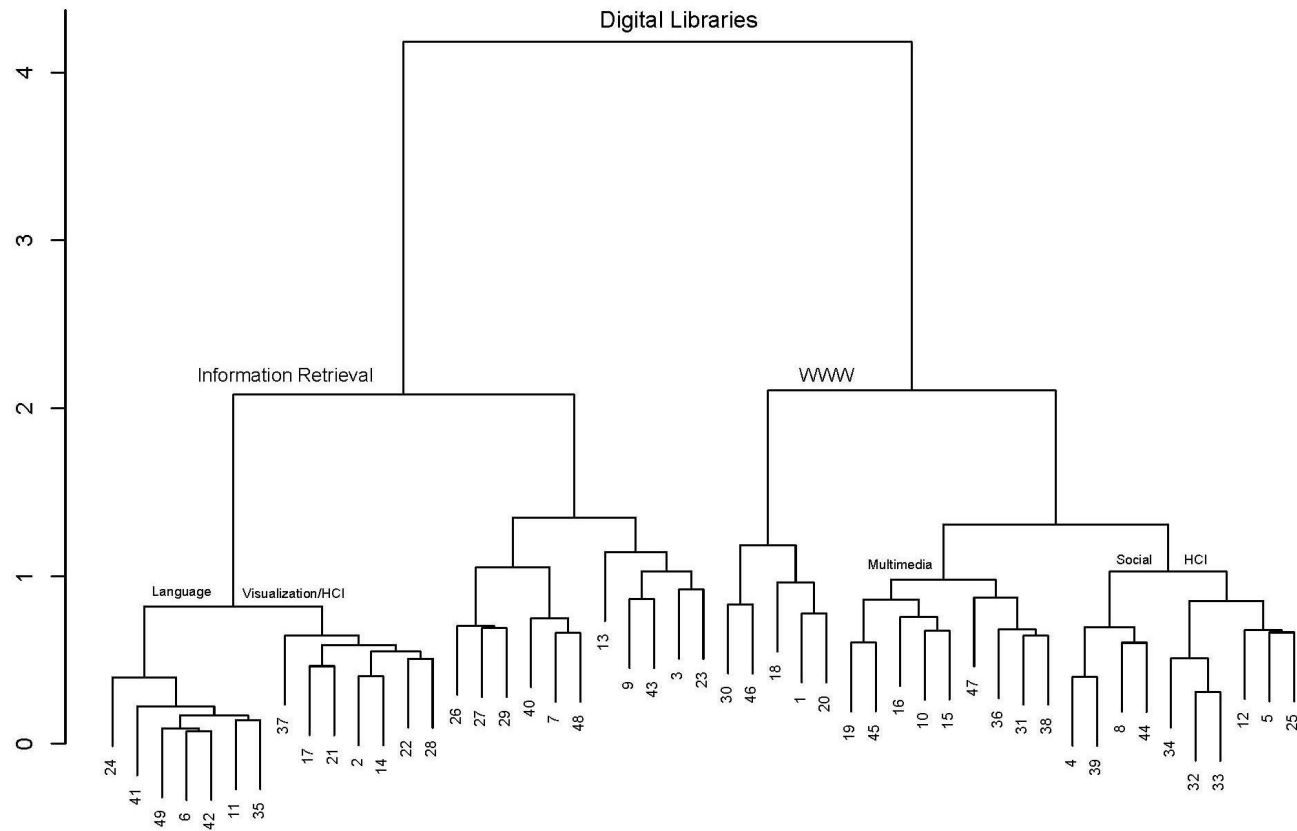
- Referees bid on papers according to this scale:
 - 0: did not provide bid data
 - 1: expert in domain and wants to review paper
 - 2: expert in domain and doesn't care to review paper
 - 3: non-expert
 - 4: conflict of interest

Sub	Ref1	Ref2	Ref3	Ref4
15	4	2	2	1
16	3	3	2	1
17	0	2	1	4
18	2	2	3	3

Optimal Referee Inclusion Value

$$0.0 = \mathbf{G4} \sim \mathbf{G3} \ll \mathbf{G2} \sim \mathbf{G1} = 1.0$$





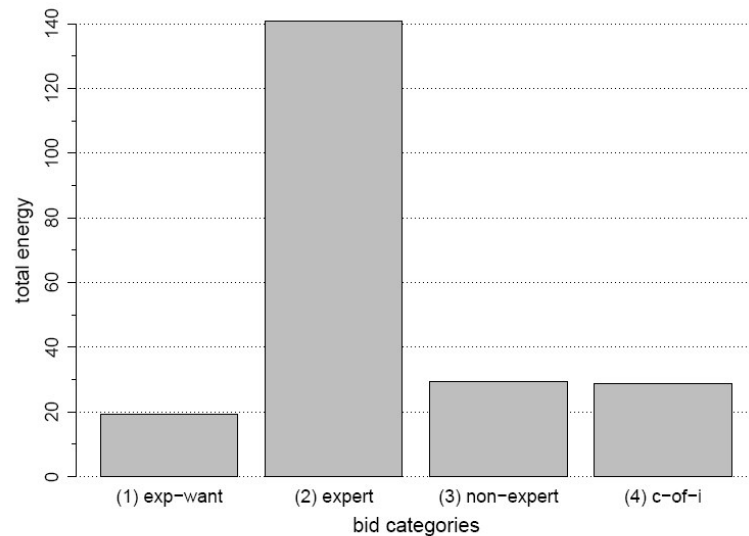
Co-Authorship Relative Rank vs. Referee Similarity Matrix

Degrees of freedom = 2399
 $p < 2.2^{-16}$
Pearson Correlation of **0.383**

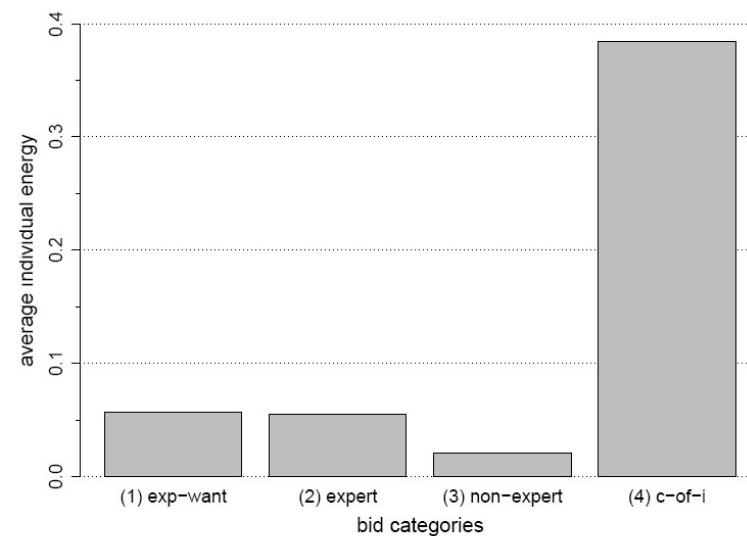
- Therefore DBLP Co-Authorship Network is correlated with the bidding behavior of the referee similarity matrix.
- Both represent a similar aspect of the scientific community: namely the relative expertise of scientists.

Rodriguez, M.A., Bollen, J., "Simulating Network Influence Algorithms Using Particle-Swarms: PageRank and PageRank-Priors", *[submitted]*, September 2005.

Results of a Energy Distribution within the DBLP Co-Authorship Network



Total Energy for each Group

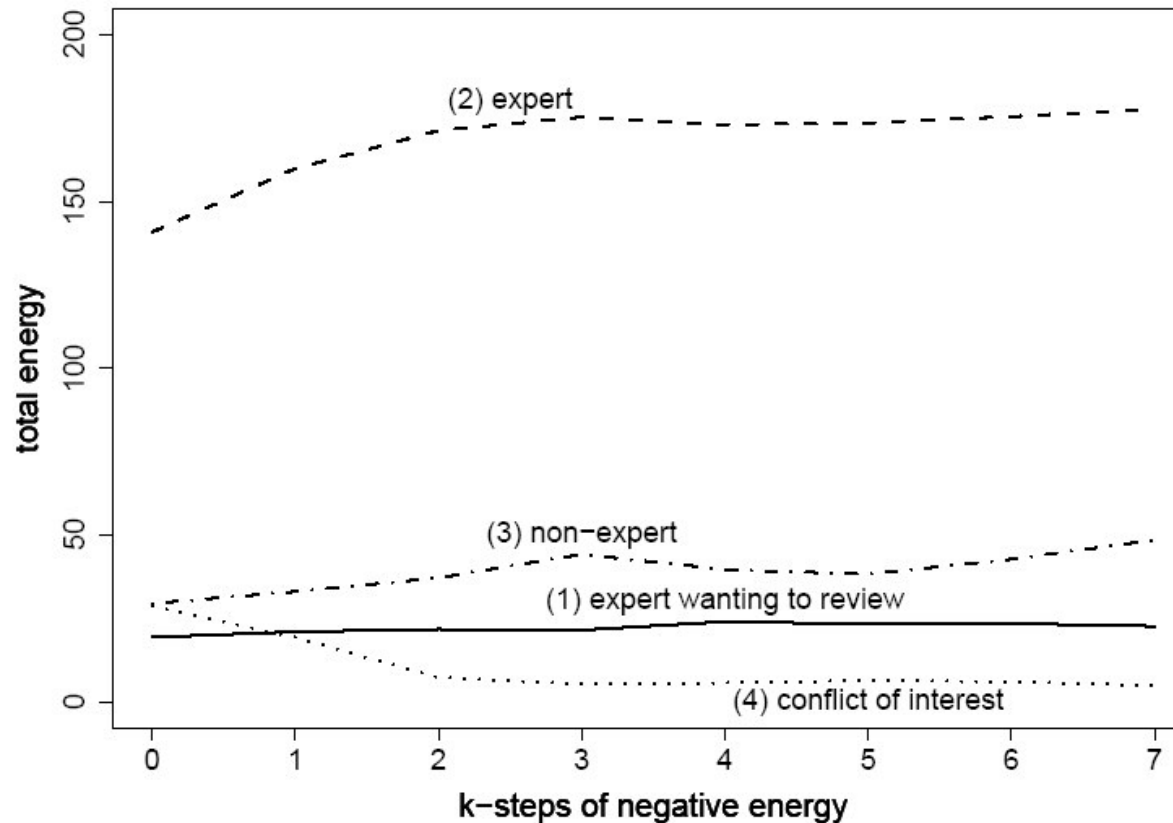


Normalized by population

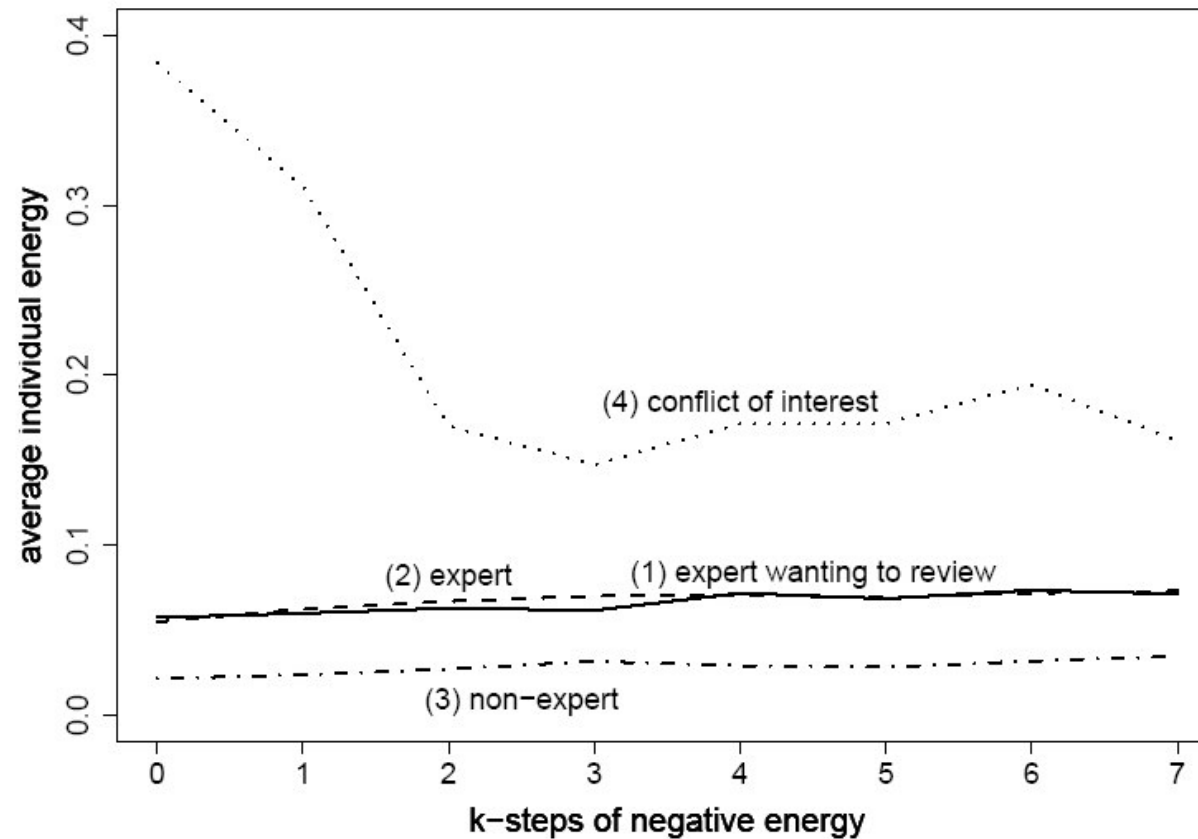
Automatic Solicitation of Referees



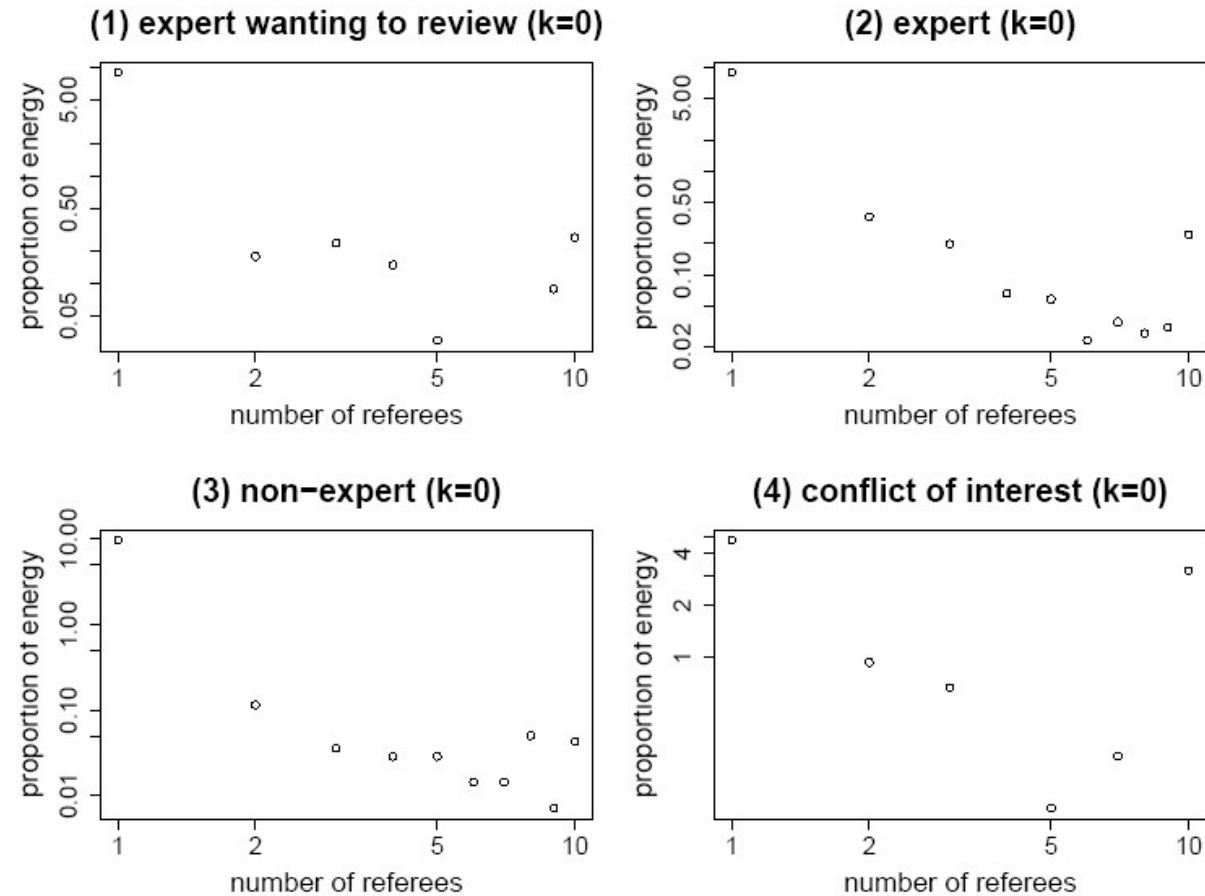
The Inclusion of Negative Energy to Curtail Conflict-of-Interest Scenarios



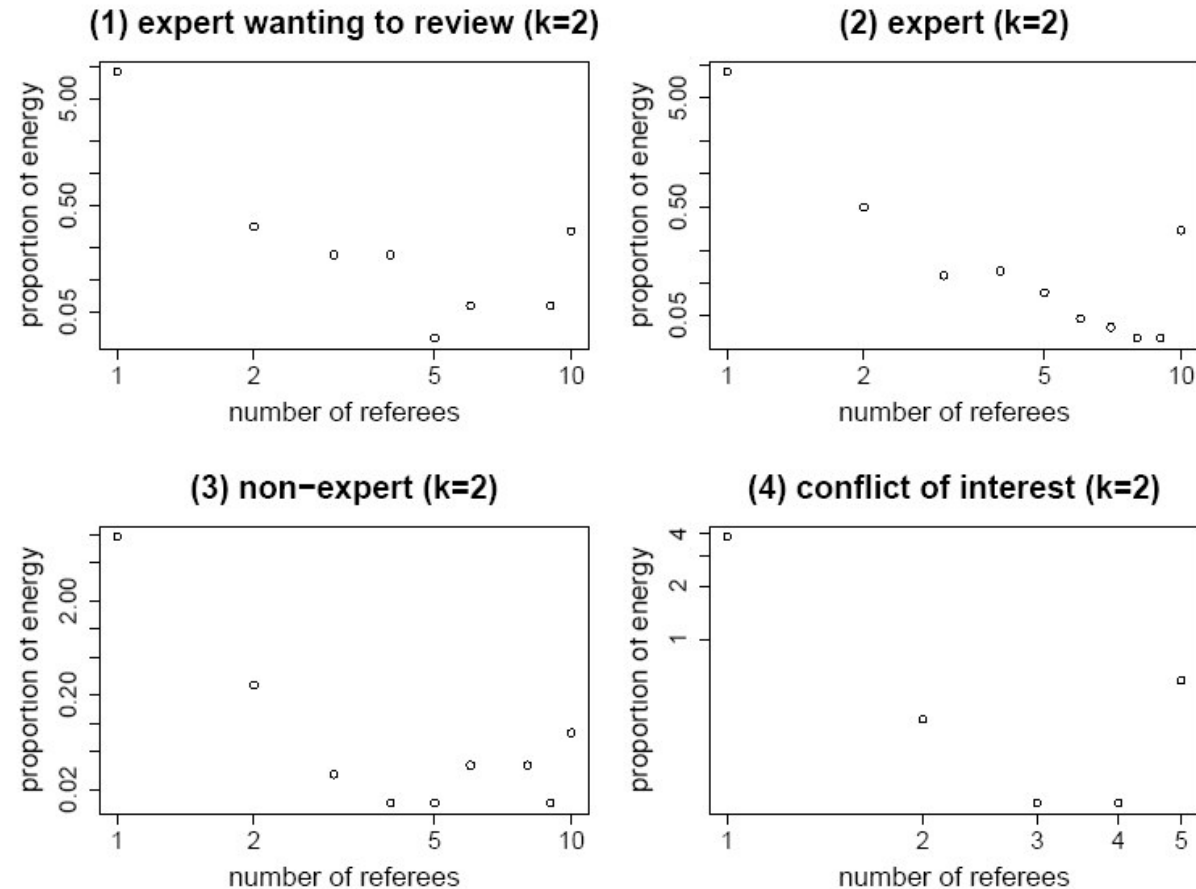
The Inclusion of Negative Energy to Curtail Conflict-of-Interest Scenarios



Energy Distribution Amongst the 4 Groups (k=0)

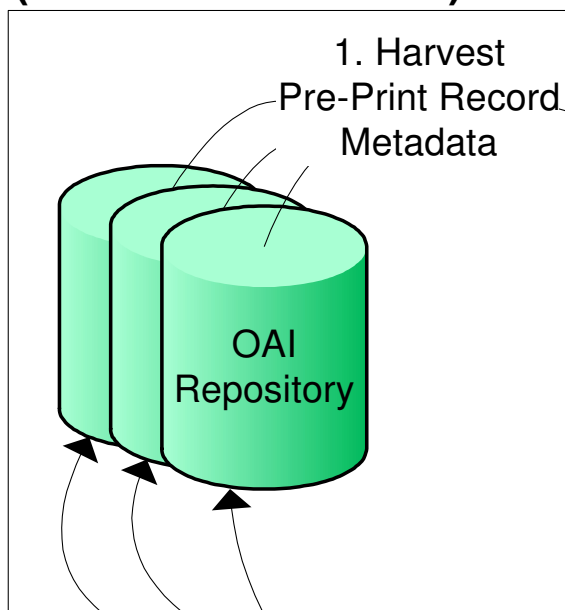


Energy Distribution Amongst the 4 Groups (k=2)

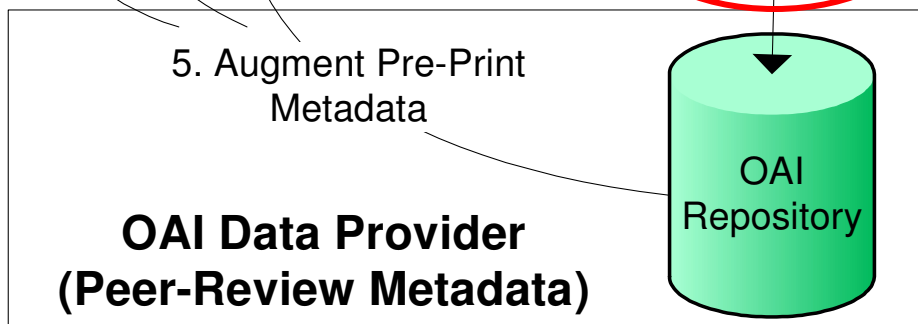
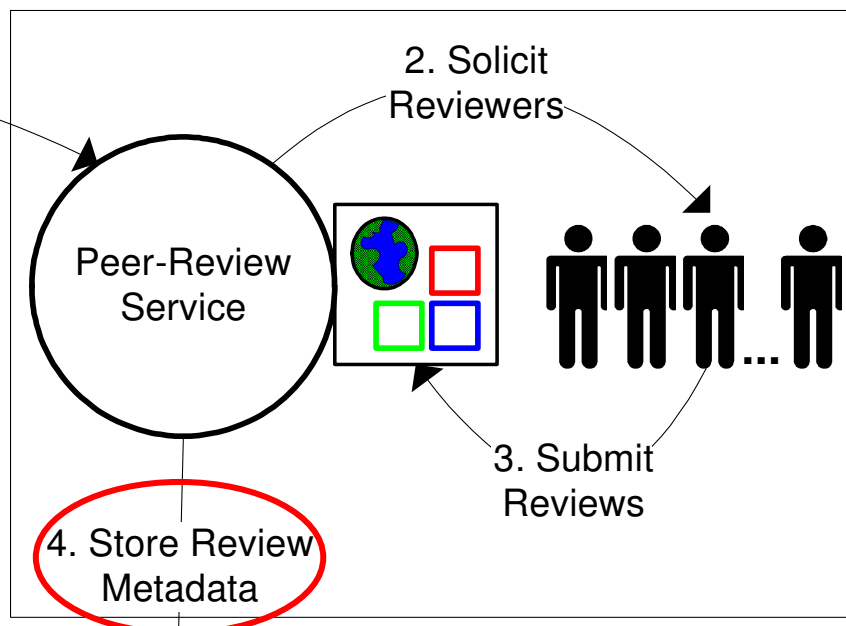


Overview of the OAI Peer-Review Architecture

OAI Data Provider (Pre-Print Metadata)



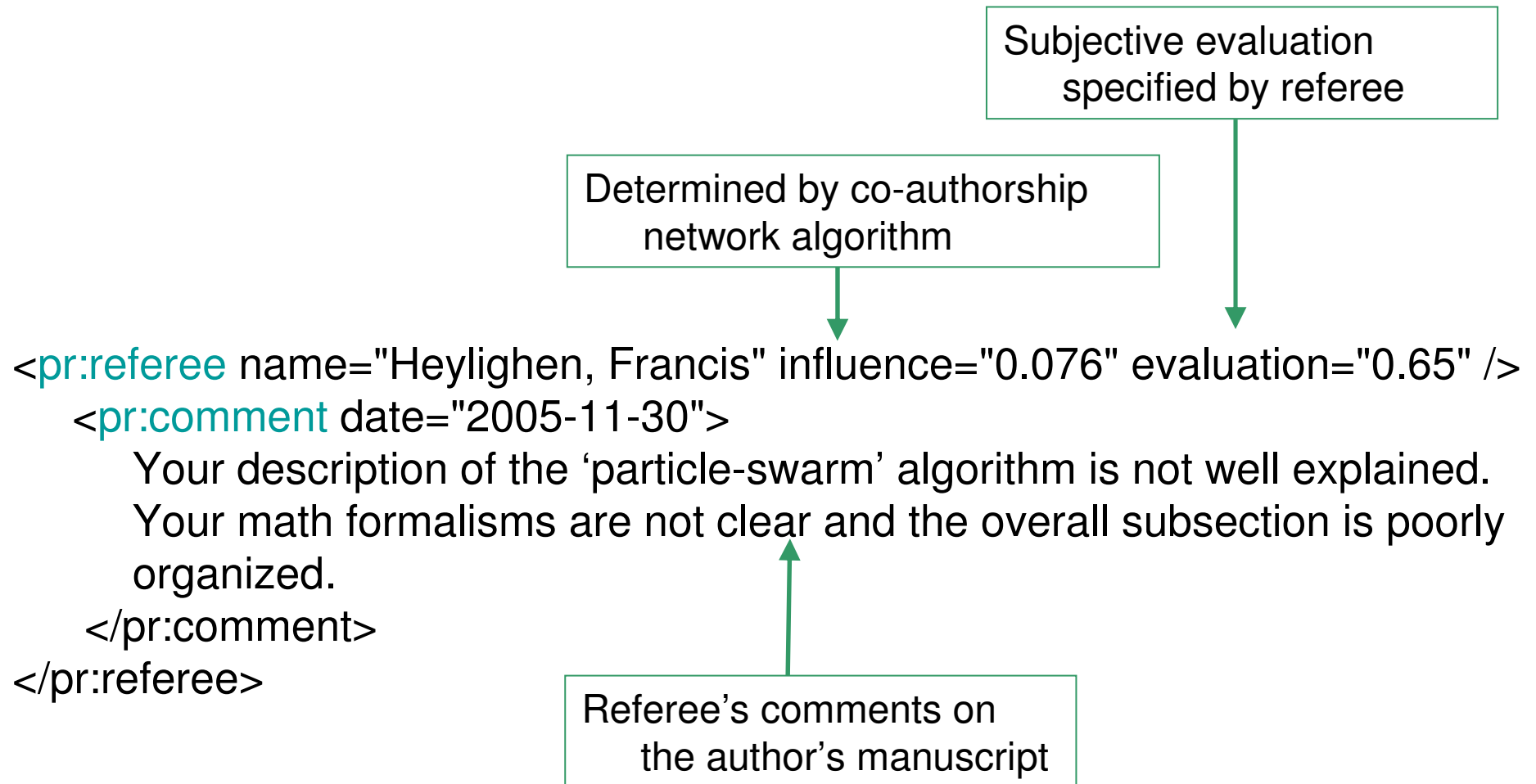
OAI Service Provider



3 Proposed Metadata Tags for the Pre-Print's Metadata Record

- <pr:review>
 - <pr:referee>
 - <pr:comment>

Peer-Review Metadata



Peer-Review Metadata

[http://peer.review.service.org/oai2?
verb=GetRecord&identifier=oai:arXiv.org:cs/0504084&metadataPrefix=pr](http://peer.review.service.org/oai2?verb=GetRecord&identifier=oai:arXiv.org:cs/0504084&metadataPrefix=pr)

```
<record>
  <header>
    <identifier>oai:arXiv.org:cs/0504084</identifier>
    <timestamp>2005-04-24</timestamp>
    <setSpec>cs</setSpec>
  </header>
  <metadata>
    <pr:review evaluation="0.755" stability="0.50">
      <pr:referee name="Heylighen, Francis" influence="0.076" evaluation="0.65" />
      <pr:comment date="2005-11-30">
        Your description of the 'particle-swarm' algorithm is not well explained.
        Your math formalisms are not clear and the overall subsection is poorly
        organized.
      </pr:comment>
    </pr:referee>
    <pr:referee>
      ...
    </pr:referee>
  </pr:review>
</metadata>
</record>
```

Evaluation and Stability Metadata Scores

-Evaluation:

$$E = \frac{\sum_{x=0}^{|A|} \inf(n_x) \cdot \text{eval}(n_x)}{\sum_{x=0}^{|A|} \inf(n_x)}$$

-Stability:

*** Simple average of the evaluations of all participating referees.**

$$S = \sum_{x=0}^{|A|} \inf(n_x)$$

*** Stability allows the community to know how much of the reviewer influence has been associated with an evaluation.**

So?

- The separation between certification and dissemination.
- Scholarly communication process solely mediated by the scholarly community. No third part intervention.
- A quantitative representation of the peer-review process. Therefore, the peer-review process can become the object of scientific inquiry.
- In combination with OAI repositories, a publication model that has limited monetary overhead.

Questions?

Refer to heading.